PROCEEDINGS OF TDWG

Biodiversity Information Science and Standards 1: e20315
doi: 10.3897/tdwgproceedings.1.20315

OPEN ACCESS

Conference Abstract

AnnoSys - an online tool for sharing annotations to enhance data quality

Okka Tschöpe‡, Lutz Suhrbier‡, Anton Güntsch‡, Walter G. Berendsohn‡

‡ Botanic Garden and Botanical Museum, Freie Universität, Berlin, Germany

Corresponding author: Okka Tschöpe (o.tschoepe@bgbm.org), Lutz Suhrbier (l.suhrbier@bgbm.org)

Received: 15 Aug 2017 | Published: 15 Aug 2017

Citation: Tschöpe O, Suhrbier L, Güntsch A, Berendsohn W (2017) AnnoSys – an online tool for sharing annotations to enhance data quality. Proceedings of TDWG 1: e20315.
https://doi.org/10.3897/tdwgproceedings.1.20315

Abstract

AnnoSys is a web-based open-source information system that enables users to correct and enrich specimen data published in data portals, thus enhancing data quality and documenting research developments over time. This brings the traditional annotation workflows for specimens to the Internet, as annotations become visible to researchers who subsequently observe the annotated specimen.

During its first phase, the AnnoSys project developed a fully functional prototype of an annotation data repository for complex and cross-linked XML-standardized data in the ABCD (Access to biological collection data Berendsohn 2007- and Darwin Core (DwC - Wieczorek et al. 2012) standards, including back-end server functionality, web services and an on-line user interface Tschoepe et al. 2013. Annotation data are stored using the Open Annotation Data Model Sanderson et al. 2013 and an RDF-database Suhrbier et al. 2017. Public access to the annotations and the corresponding copy of the original record is provided via Linked Data, REST and SPARQL web services.

AnnoSys can easyly be integrated into portals providing specimen data (see Suhrbier & al., this session). As a result, the individual specimen page then includes two links, one providing access to existing annotations stored in the AnnoSys repository, the other linking to the AnnoSys annotation Editor for annotation input. AnnoSys is now integrated into a dozen specimen portals, including the Global Biodiversity Information Facility GBIF and the

© Tschöpe O et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Global Genome Biodiversity Network GGBN. In contrast to conventional, site-based annotation systems, annotations regarding a specimen are accessible from all portals providing access to the specimen's data, independent of which portal has originally been used as a starting point for the annotation.

Apart from that, users can query the data in the AnnoSys portal or create a subscription to get notified about annotations using criteria referring to the data record. For example, a specialist for a certain family of organisms, working on a flora or fauna of a certain country, may subscribe to that family name and the country. The subscriber is notified by email about any annotations that fulfil these criteria. Other possible subscription and filter criteria include the name of collector, identifer or annotator, catalogue or accession numbers, and collection name or code. For curators a special curatorial workflow supports their handling of annotations, for example confirming a correction according to the annotation in the underlying primary database.

User feedback on the currently available system has led to a significantly simplified version of the user interface, which is currently undergoing testing and final implementation. Moreover, the current, second project phase aims at extending the generic qualities of AnnoSys to allow processing of additional data formats, including RDF data with machine readable semantic concepts, and thus opening up the data gathered through AnnoSys for the Semantic Web. We developed a semantic concept driven annotation management, including the specification of a selector concept for RDF data and a repository for original records extended to RDF and other formats. Based on DwC RDF terms and the ABCD ontology, which deconstructs the ABCD XML-schema into individually addressable RDF-resources, we built an “AnnoSys ontology”.

The AnnoSys-2 system is currently in the testing phase and will be released in 2018. In future research (see Suhrbier, this volume), we will examine the use of AnnoSys for taxon-level data as well as its integration with image annotation systems. BGBM Berlin is committed to sustain AnnoSys beyond the financed project phase.

Keywords

annotation, AnnoSys, biological collection data, RDF data, semantic web, semantic concept, ontology, open kowledge

Presenting author

Okka Tschöpe & Walter G. Berendsohn

Funding program

AnnoSys-2 is funded by the German Research Foundation (DFG, BE 2283/4-2).

Hosting institution

Botanic Garden and Botanical Museum, Freie Universität Berlin

References

- Berendsohn WG (2007) Access to biological collection data. ABCD Schema 2.06 – ratified TDWG Standard. Berlin: Botanischer Garten und Botanisches Museum Berlin-Dahlem (BGBM), Freie Universität Berlin. http://www.bgbm.org/TDWG/CODATA/Schema/default.htm. Accessed on: 2017-8-15.
- Sanderson R, Ciccarese P, Van de Sompel H (2013) W3C Open Annotation Data Model, Community Draft, 08 February 2013. http://www.openannotation.org/spec/core/. Accessed on: 2017-8-14.
- Suhrbier L, Kusber W-, Tschöpe O, Güntsch A, Berendsohn WG (2017) AnnoSys—implementation of a generic annotation system for schema-based data using the example of biodiversity collection data. Database 2017 (1): . https://doi.org/10.1093/database/bax018
- Tschoepe O, Macklin J, Morris R, Suhrbier L, Berendsohn W (2013) Annotating biodiversity data via the Internet. Taxon 62: 1248-1258. https://doi.org/10.12705/626.4
- Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D (2012) Darwin Core: an evolving community-developed biodiversity data standard. PloS one 7 (1): e29715. https://doi.org/10.1371/journal.pone.0029715